Listening without hearing

Nadia M. Biassou MD, PhD
Senior Research Physician
Dept of Radiology
Division of Neuroradiology
NIH Clinical Center
and
Senior Fellow
Linguistics Data Consortium
Graduate School of Arts and Sciences
University of Pennsylvania

Financial Disclosure

- None

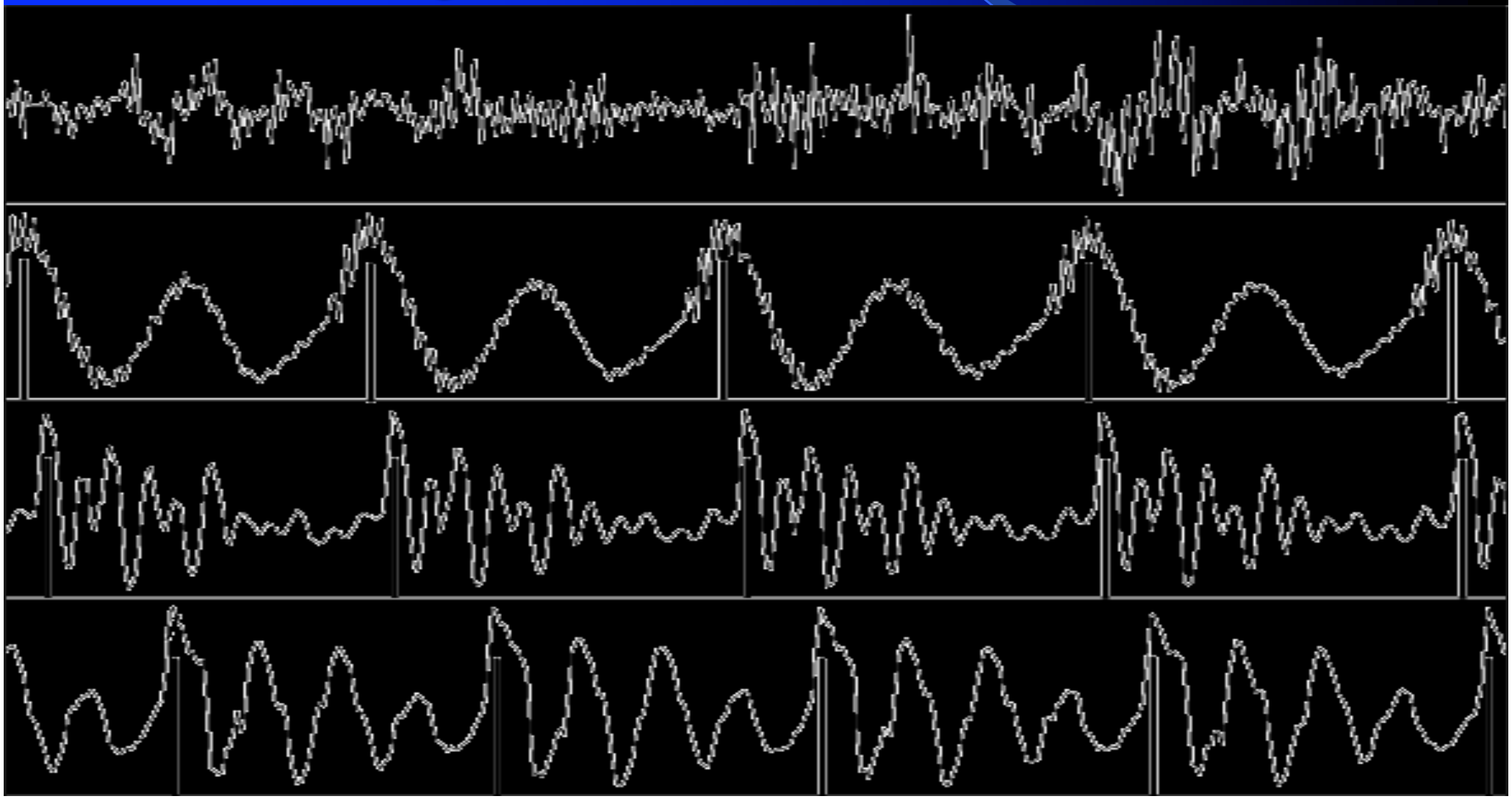
Learning Objectives

- The basics of the physics of speech
- What is currently known about conscious neurobiologic speech perception?
- Can unconscious speech perception by reliably measured?
- What can its study tell us about the general nature of speech perception and about the human brain that processes it?

Introduction

- What is speech and why is it special?

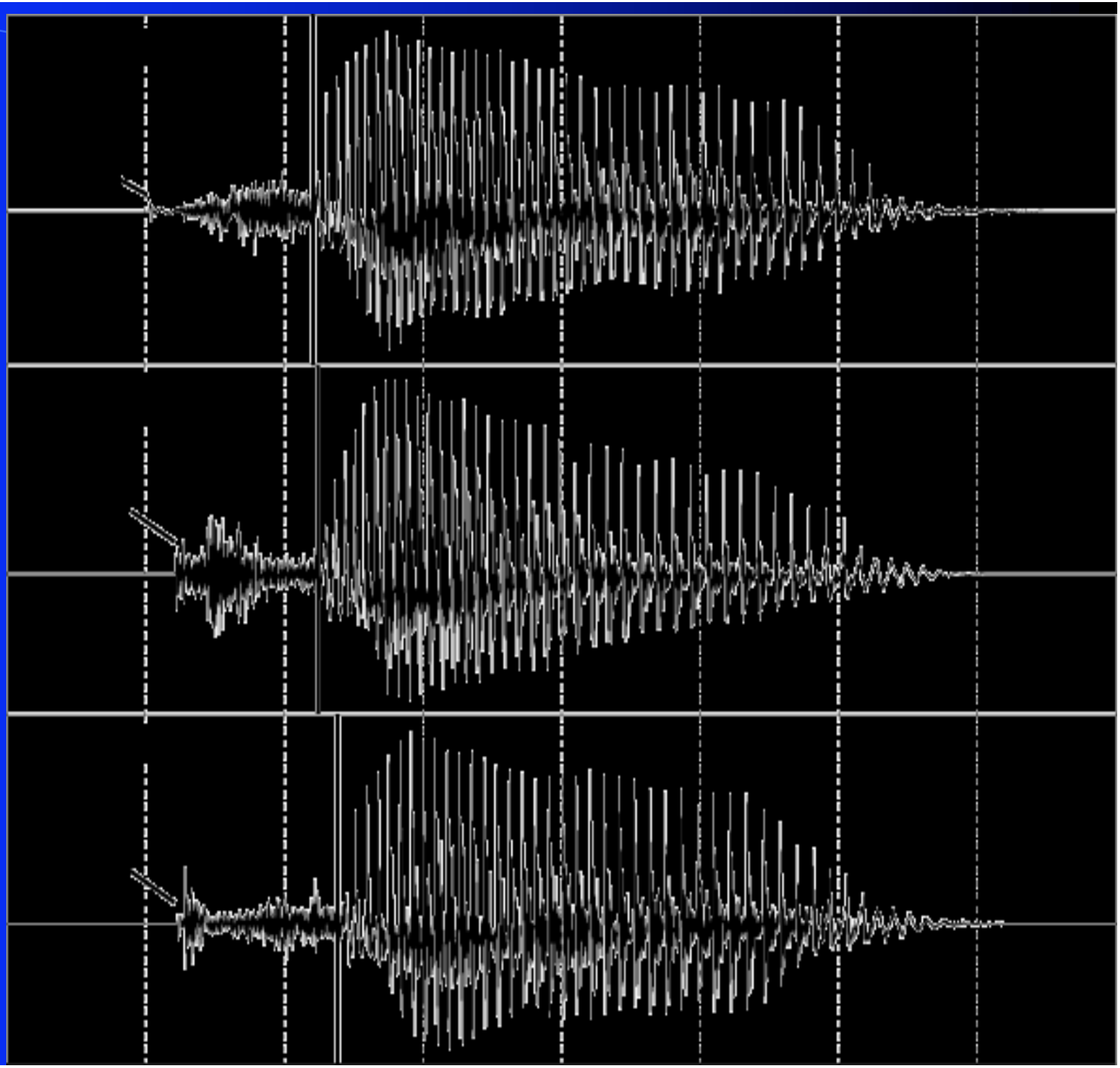
Speech is the entryway to human linguistic communication

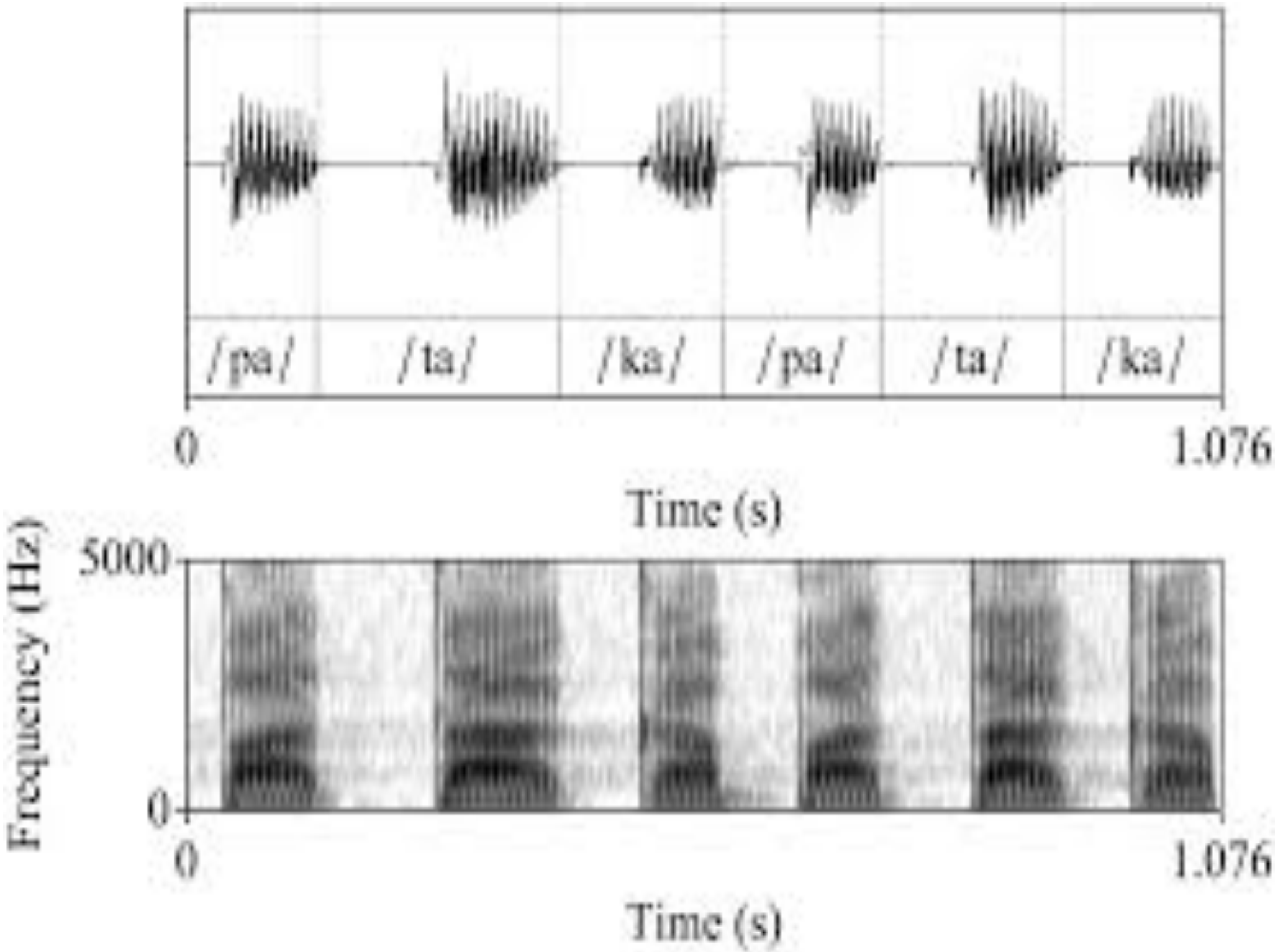


1 Pa

2. Ta

3. Ka

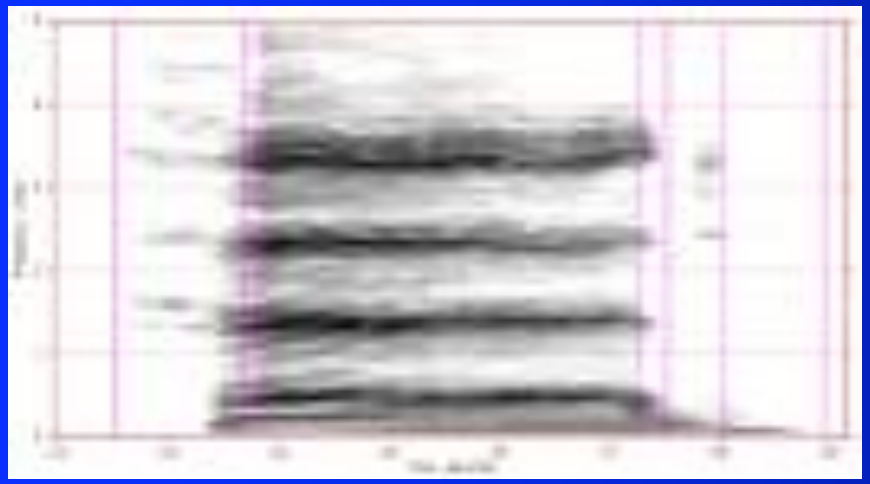



/pa/
/ta/
/ka/
/pa/
/ta/
/ka/
0
1.076
Time (s)
5000
Frequency (Hz)
0
0
1.076
Time (s)

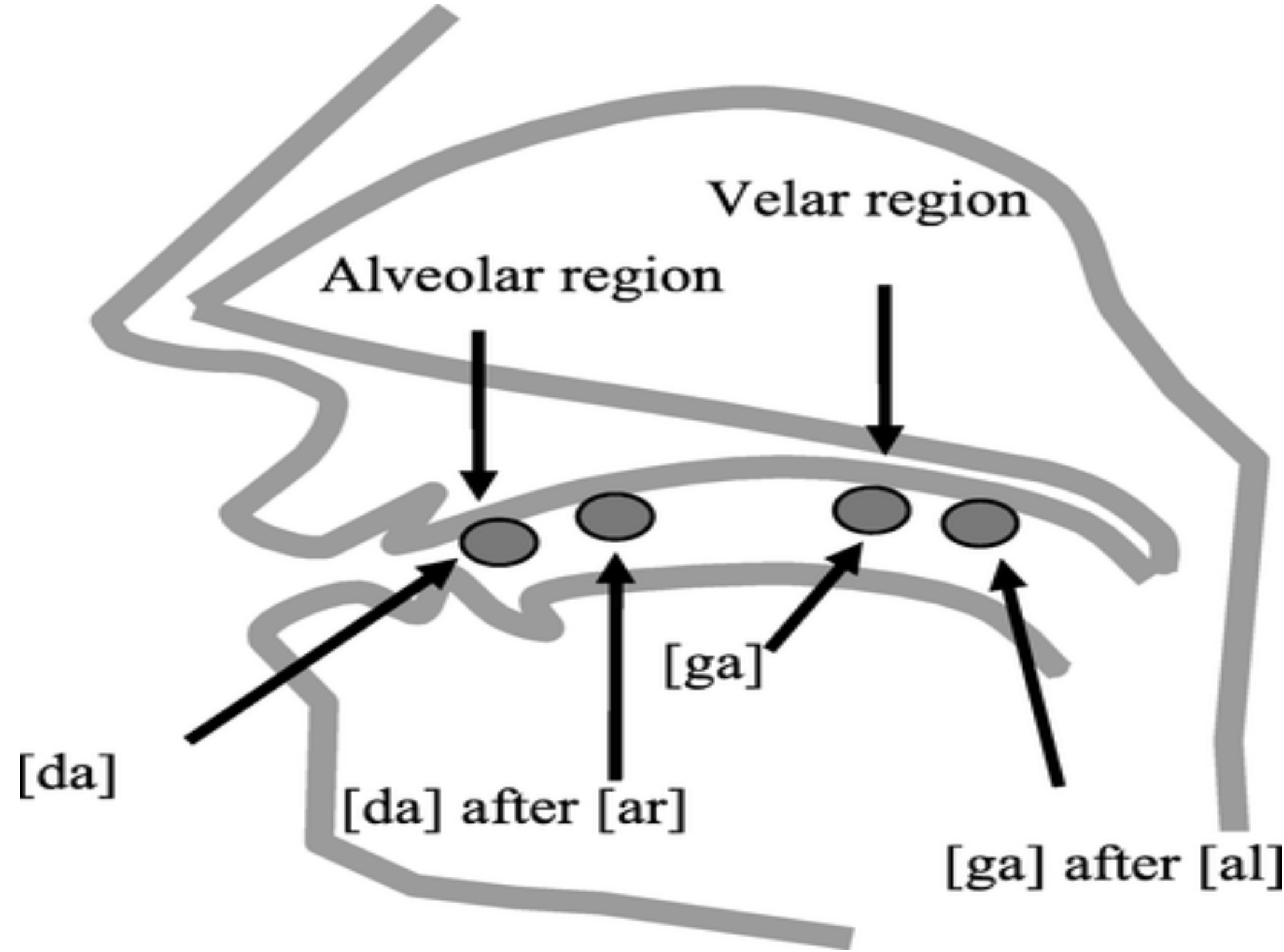
Formant frequency

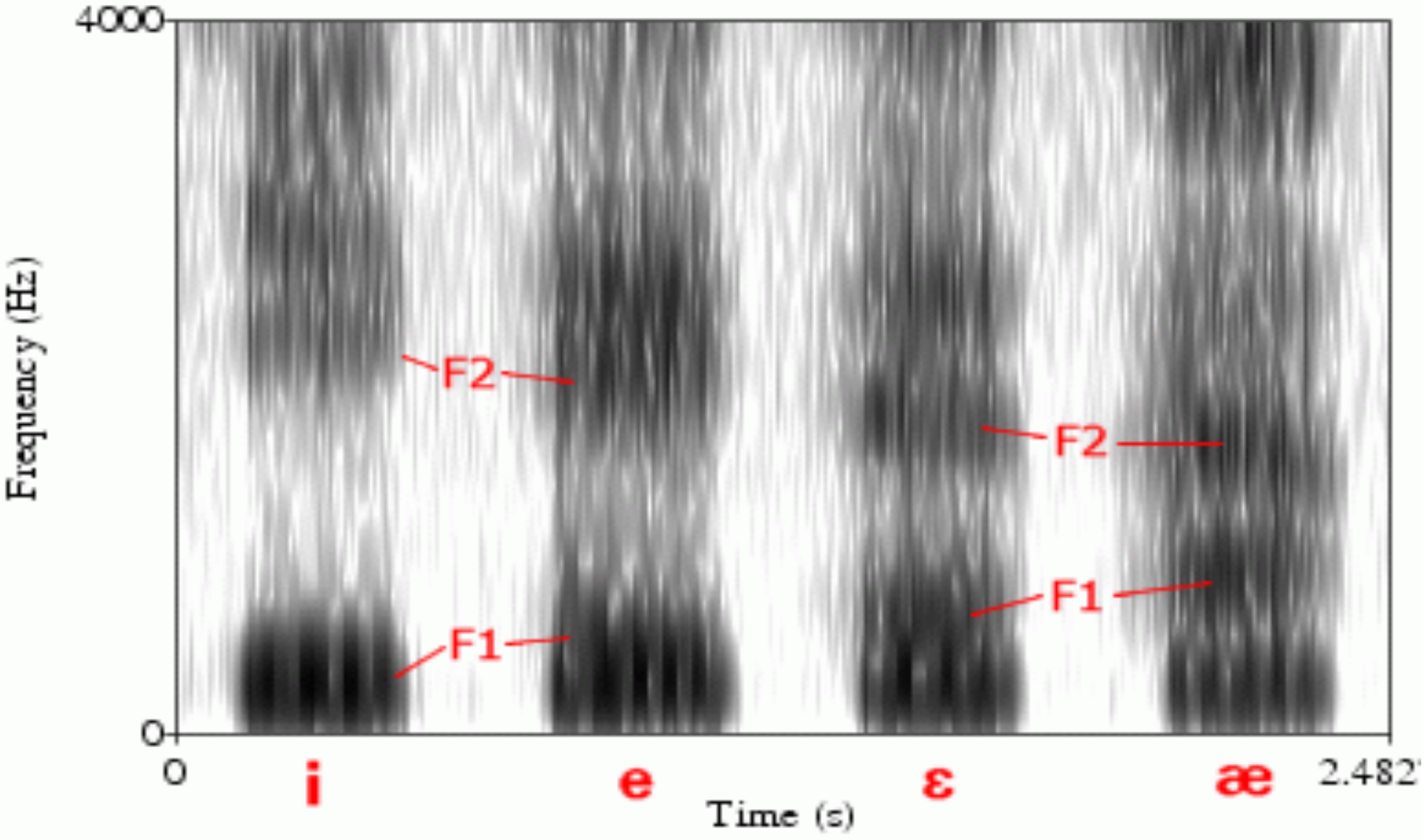
- F0 is called the fundamental frequency and represents the frequency of vocal cord oscillation

Formant frequencies



- Oscillation of vocal cords and its harmonics
- F0 1
- F1 3
- F2 5
- F3 7

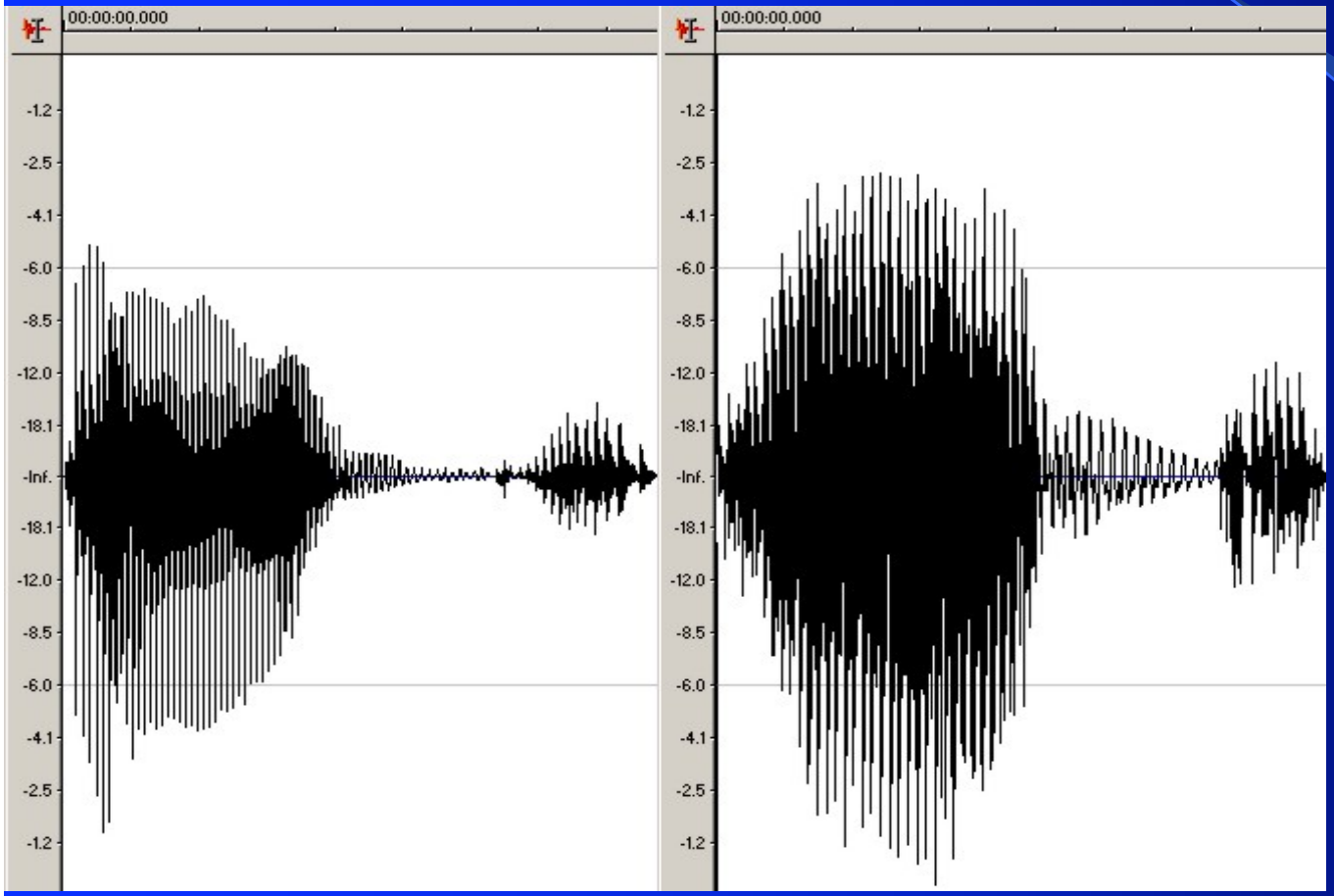

Velar region
Alveolar region
[ga]
[da]
[da] after [ar]
[ga] after [al]


4000
Frequency (Hz)
F2
F2
F1
F1
0
0
i
e
ɛ
æ
2.482
Time (s)

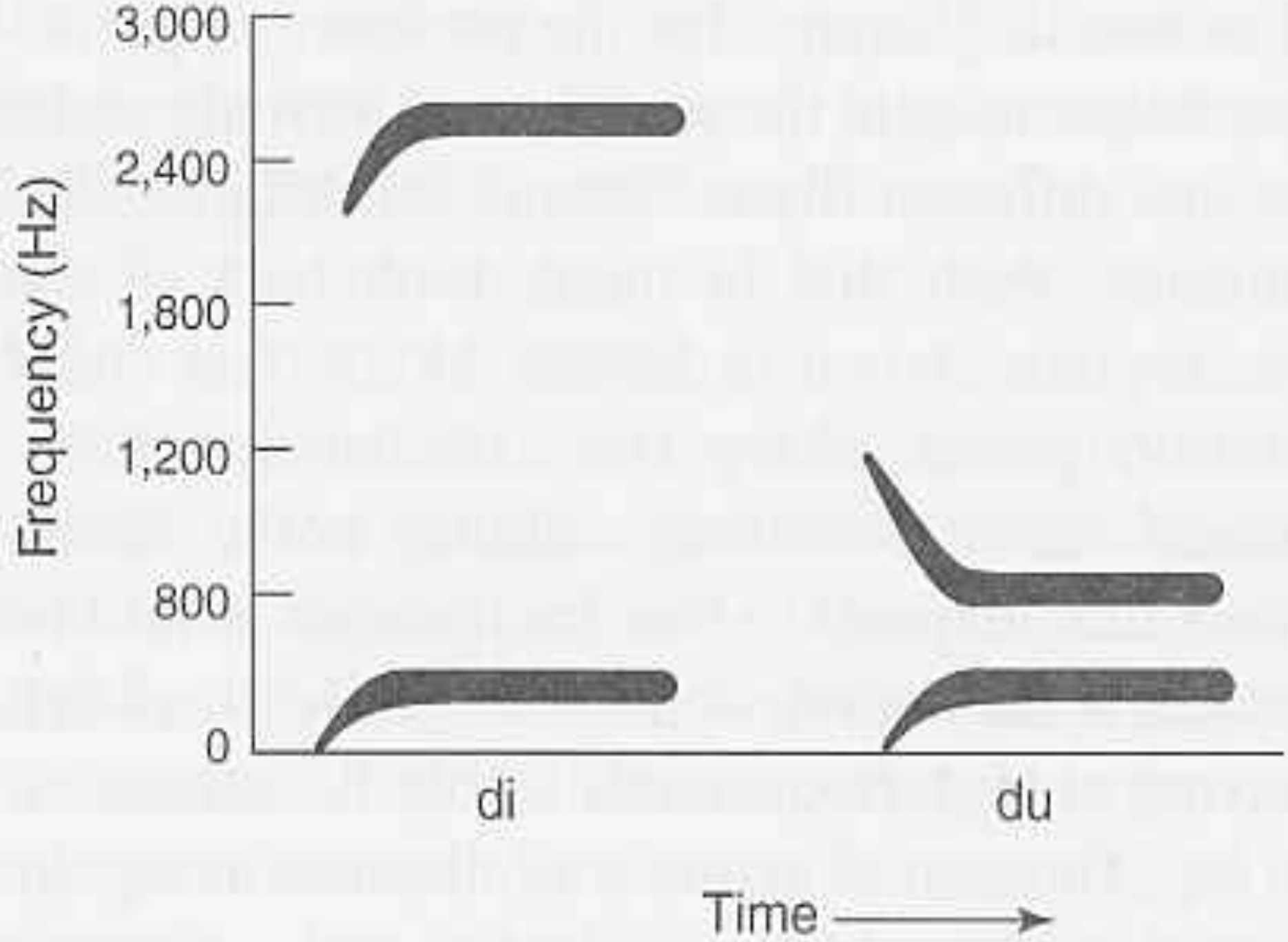
The speech waveform

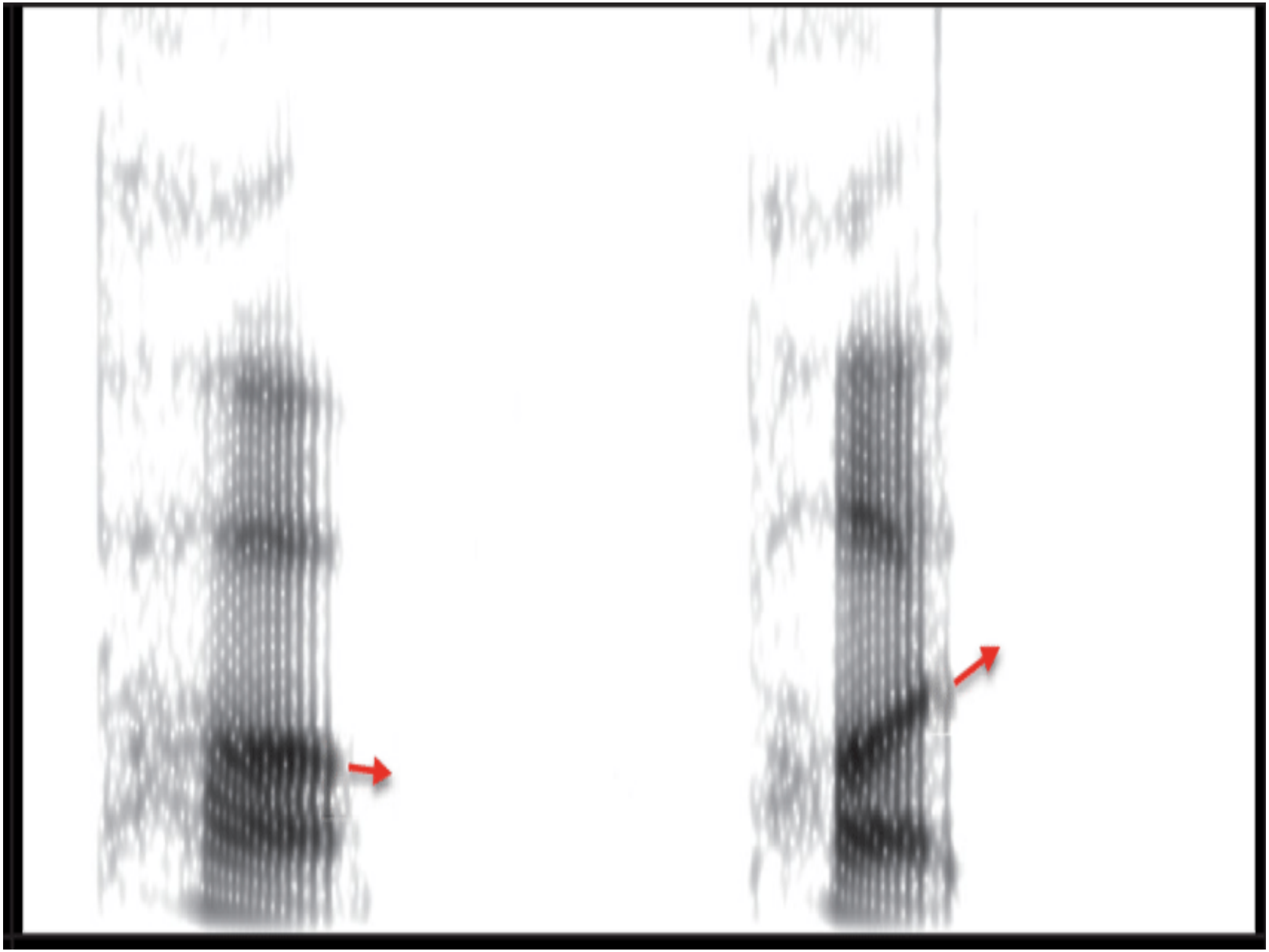
- The production of any sound during word production is simultaneously influenced by the sounds that precede and follow it.
- Liberman et al., 1957

Coarticulation of sounds

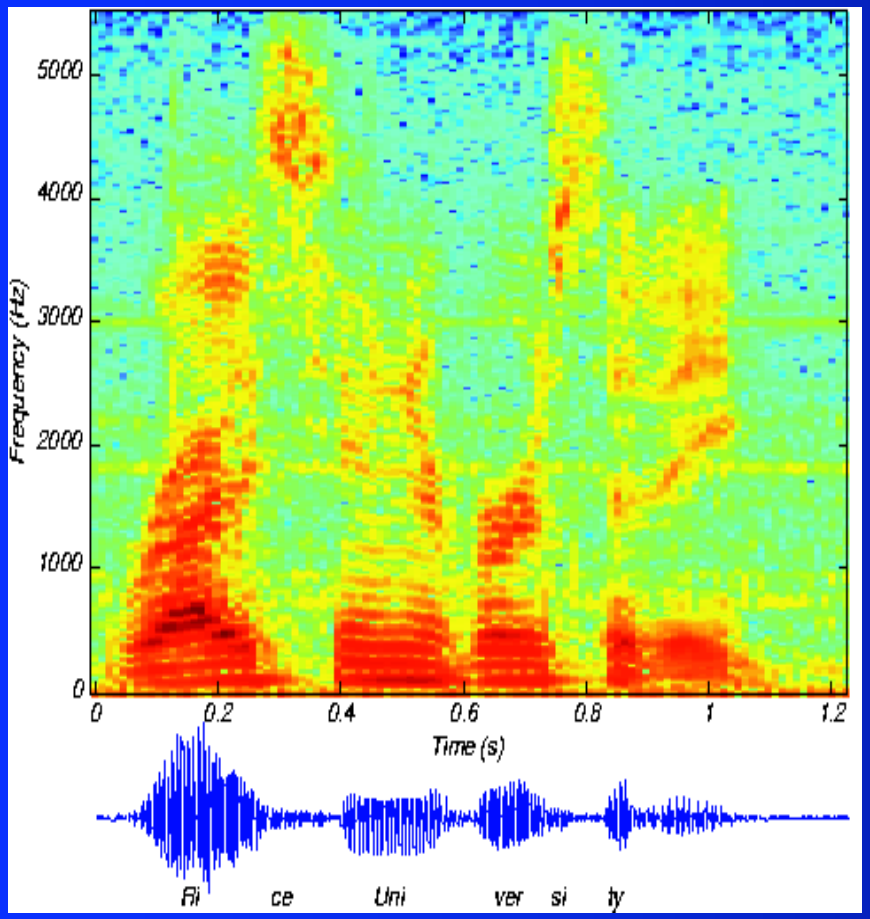


“ebb” vs. “egg”


3,000
2,400
1,800
1,200
800
0
Frequency (Hz)
di
du
Time



The speech spectrograms: formant frequency transitions



- The formant frequencies transitions reflect coarticulation

Does the brain listen to every acoustic variation during speech perception?

Bottom Up processes

Bottom-up processing refers to processing sensory information as it is coming in

TASK

- PART 1: Actively decided whether real and nonreal words are real words of English, half of the real and nonreal words are acoustically manipulated

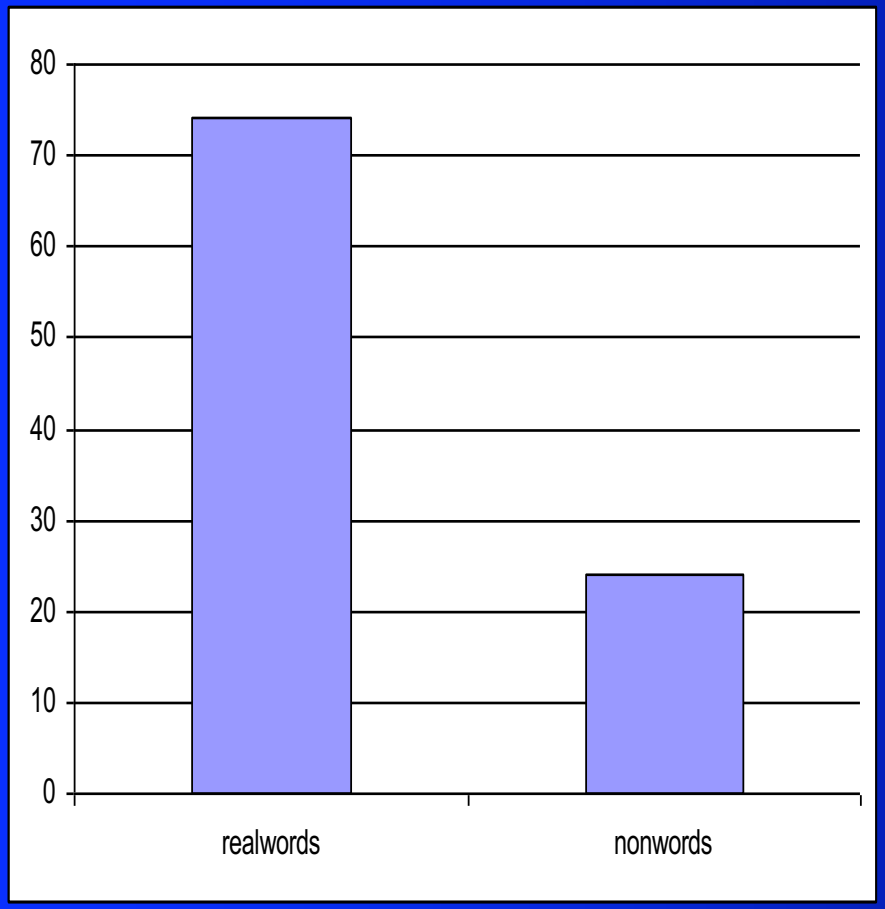
STIMULI

- EXPERIMENT 1

- 40 REAL WORDS
  - HALF ARE ACOUSTICALLY MANIPUALTED
  - HALF ARE NON-MANIPULATED

- 40 NONREAL WORDS
  - HALF ARE ACOUSTICALLY MANIPULATED
  - HALF ARE NON-MANIPULATED

RESULTS



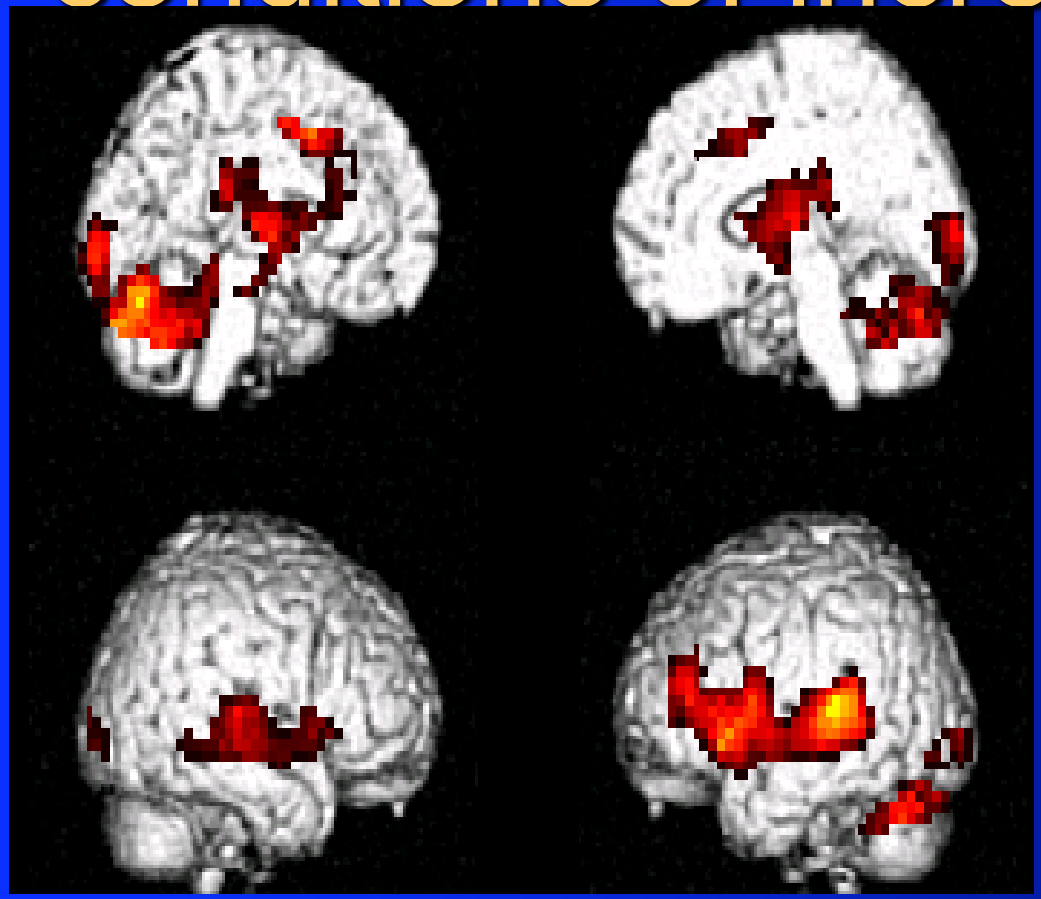
- The brain takes 74msecs longer to process the acoustically manipulated realwords, even though subjects could not consciously distinguish the word types

Sensory changes affect higher order language processing

BLUMSTEIN and colleagues

- LEXICAL DECISION TASKS IN WHICH LEXICAL ITEMS WERE MANIPULATED ACOUSTIC GAP DETECTION (I.E. VOT) below the conscious level
- Sensory alteration can affect activation semantic priming and lexical access.
- FARAH et al argue that words may also be stored with visual associated information.

What are the neural networks that subserve subconscious processing of speech during conditions of increased effort?



- LEFT INFERIOR FRONTAL CORTEX, ANTERIOR CINGULATE AND THALAMUS
- POSTERIOR SUPERIOR TEMPORAL LOBES BILATERALLY
- OCCIPITAL LOBES
- LEFT CEREBELLUM

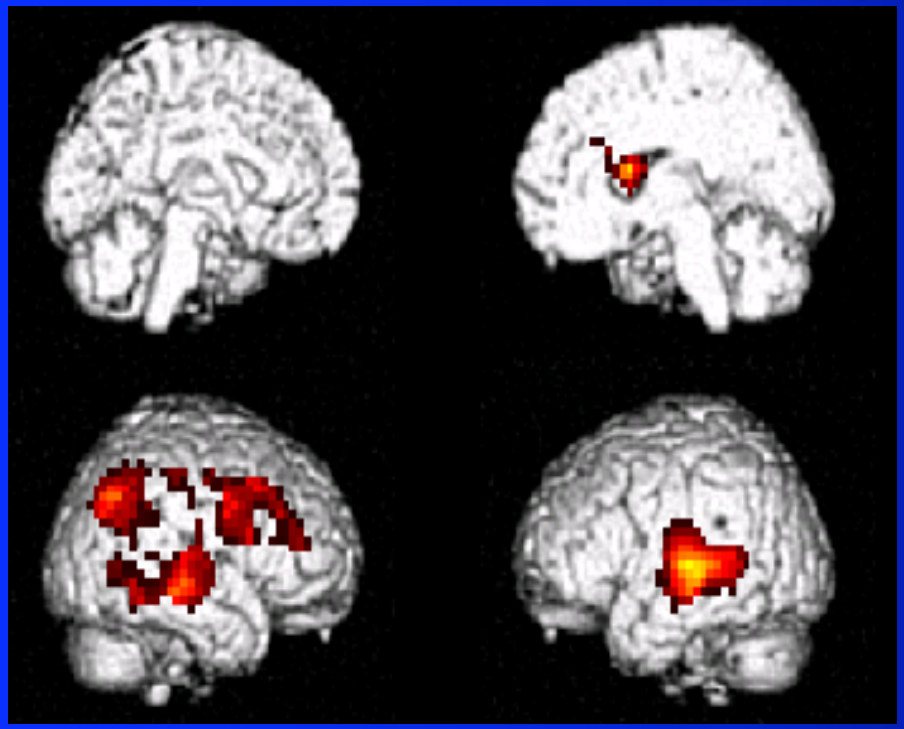
PART 2: Passively listen to real and nonreal words of English, half of which had been acoustically manipulated.

STIMULI- PARTS 1 & 2 ARE MATCHED IN WORD FREQUENCY, WORD LENGTH, NUMBER OF SYLLABLES, AND IMAGEABILITY

- EXPERIMENT 2
- 40 REAL WORDS
- HALF ARE ACOUSTICALLY MANIPULATED
- HALF ARE NOT MANIPULATED

- EXPERIMENT 2
- 40 NONREAL WORDS
- HALF ARE ACOUSTICALLY MANIPULATED
- HALF ARE NOT MANIPULATED

Are the same networks activated in conditions of less effort?



- Activation in (b) posterior superior temporal lobes and anterior cingulate are sufficiently *robust* even for the passive presentation of subconsciously manipulated realwords. But *right frontal* and *right parietal lobe networks* are activated

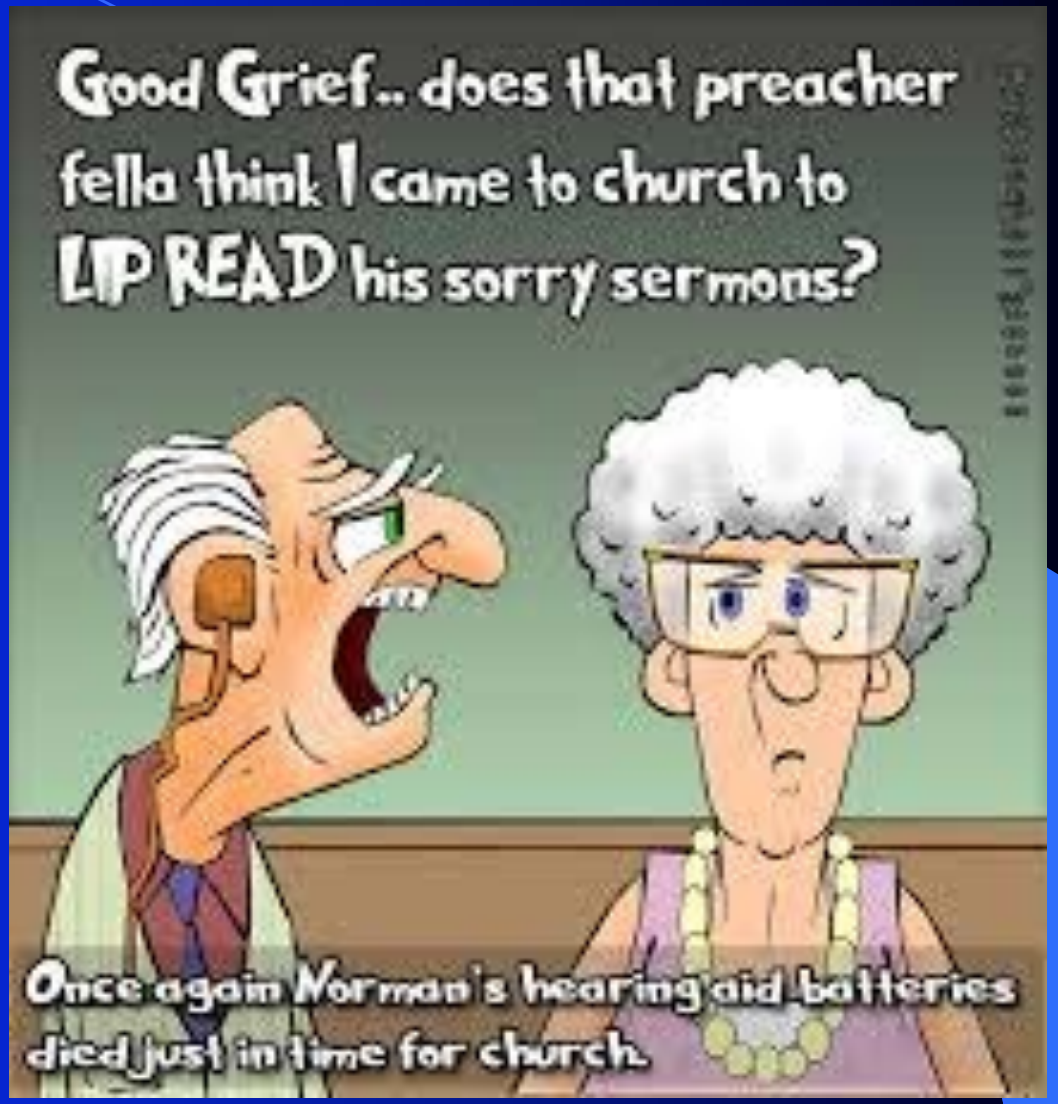
BUT IS SPEECH PERCEPTION ALL BOTTOM UP?

Top Down Processes

Visual Cues and Speech Perception

McGuck Effect

Baysan, U. (July 2017) "McGurk Effect" in F. Macpherson (ed.), The Illusions Index. Retrieved from https://www.illusionsindex.org/i/mcgurk-effect.



McGuck Effect

BBC – Horizon: Is Seeing Believing Nov 2010

TOP DOWN PROCESSES

CONTEXT AND SPEECH PERCEPTION

PHONEME RESTORATION (Warren & Warren 1970)

Phoneme Restoration Effect

- Warren & Warren (1970)
  - *It was found that the *eel was on the axle*
  - *It was found that the *eel was on the shoe*
  - *It was found that the *eel was on the orange*
  - *It was found that the *eel was on the table*
- * was a cough but it was heard as the missing phoneme implied by the context

Phoneme Restoration Effect

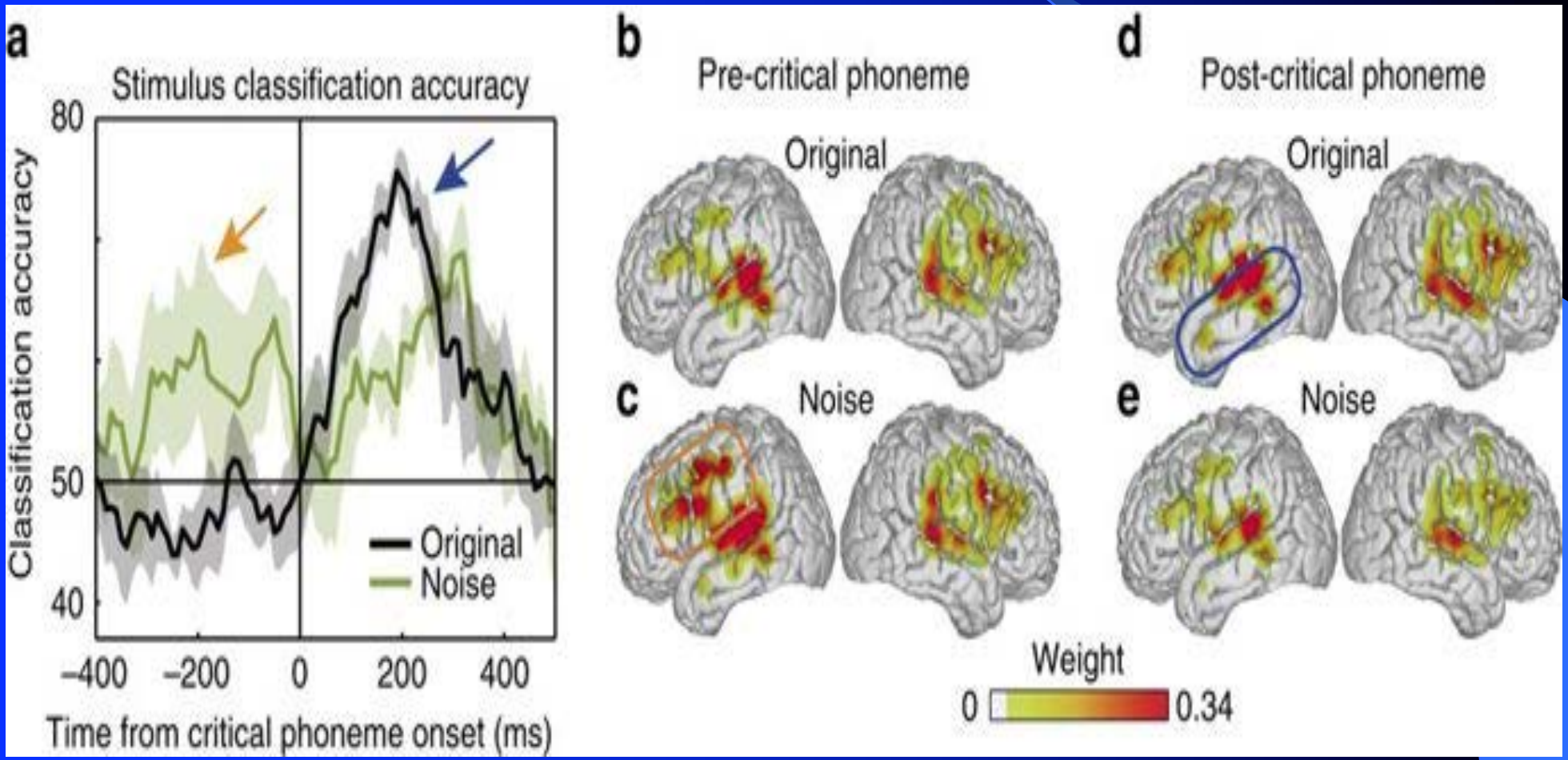
Suboptimal environment input is overridden by context of speech to hear stimuli that is in fact absent.

“The State Governors met with their respective legislatures convening in the capitol city.”

“The State Governors met with their respective le…latures convening in the capitol city.”

Ed Chang and colleagues

Leonard et al 2016, Nature Communications, 7:13619



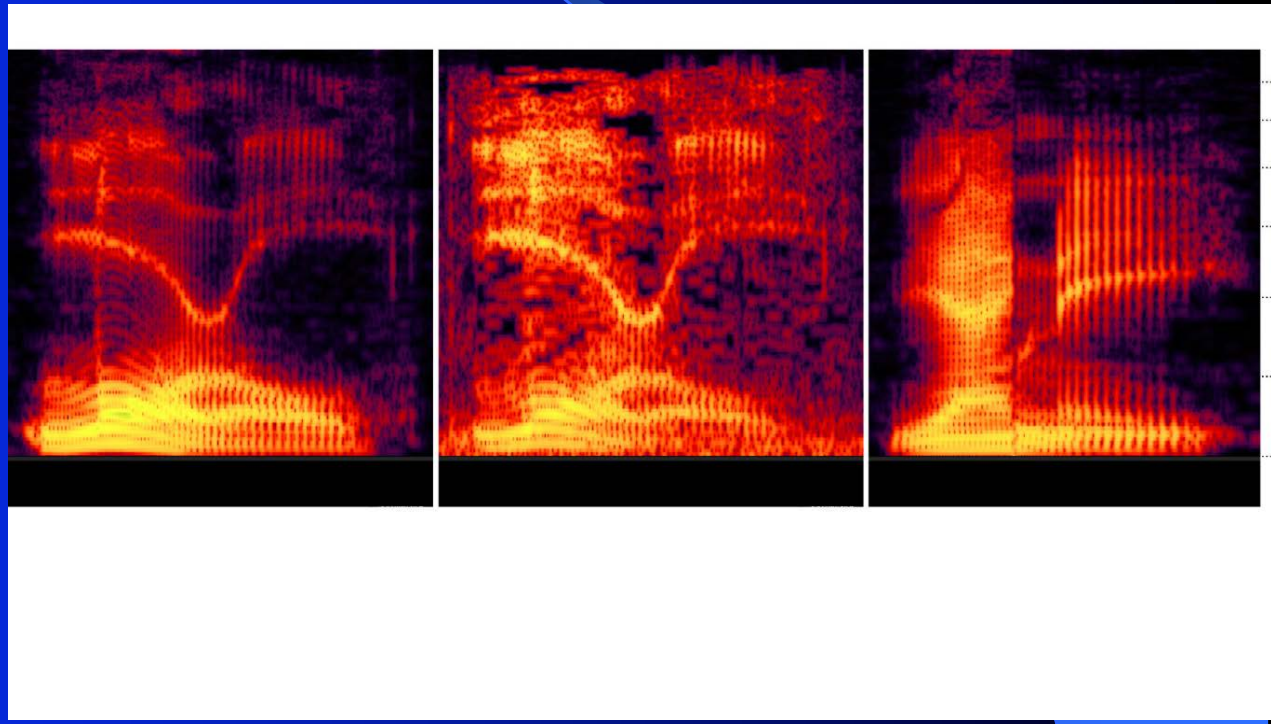
Top Down Processes

Yanny vs Laurel

Left: YANNY

Right: LAUREL

Middle spectrogram is a simulated ambiguous spectrogram – BUT listeners hear Yanny or Laurel



PERCEPTION is the point of contact between multisensory information:

BOTTOM UP (Objective)

- Processing of sensory input
- Can affect higher order cognitive and linguistic processes such as vision and semantics.

TOP DOWN (Subjective)

- Visual Input
- Context
- Linguistic phonotactics (the language that you speak) can all affect the interpretation of sensory cues.

This point of contact is dynamic in time and in space

- Different neural networks can process the same types of speech cues depending on the conditions under which the cues are being processed.
- Neural networks involved in processing subconscious fine grain speech cues can involve the right hemisphere under passive listening (or lighter attentional load)

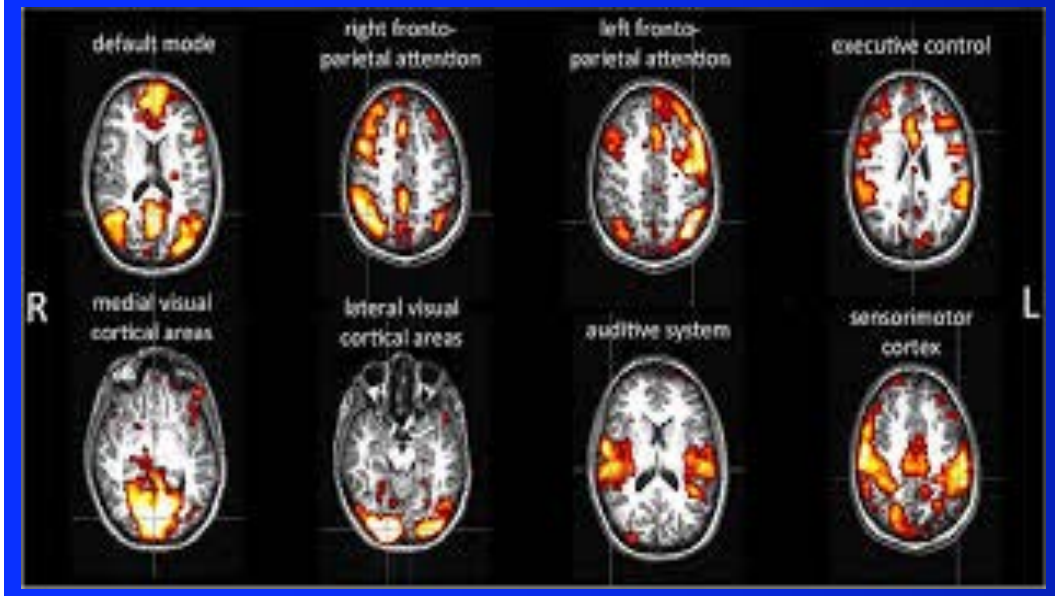
Attentional networks are always being recruited to varying degrees?

- Even for the passive listening of speech cues.

We hear want we want to hear

This doesn't only apply to cats or dogs!

Resting State fMRI



- Passive neural networks may not be fully representative of the neural networks that subserve linguistic/cognitive processes because the network dynamics change depending on the attentional load to achieve the task at hand. It is NOT solely driven by the stimulus.

CLINICAL IMPLICATIONS

- Language mapping for neurosurgery should reflect *natural state* of language processing as closely as possible including masked stimuli

New research

- Normal aging
- Alzheimer's and other neurodegenerative disorders
- Autism
- Is there a genetic basis for the balance between objective and subjective speech perception?

Future research

- Can we develop new wearable technologies that can diagnosis changes in the processing of sensory input in the preclinical stage of disease?

- “Our imaginations are limited by the knowledge that we currently possess”
  - Helen Neville (IRCS Talk, University of Pennsylvania, 1995)

- THANK YOU!